WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

INTERNATIONAL SURVEY OF THE NEURAL NETWORK RESEARCH COMMUNITY

Preliminary Report

Michael A. Rappa and Koenraad Debackere
Principal Investigators

May 25, 1990 WP# 3170-90-BPS

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge, Massachusetts 02139

INTERNATIONAL SURVEY OF THE NEURAL NETWORK RESEARCH COMMUNITY:

Preliminary Report

Michael A. Rappa and Koenraad Debackere
Principal Investigators

May 25, 1990 WP# 3170-90-BPS

We are extremely grateful to the more than 700 neural network researchers from thirty countries who participated in this survey. Their willingness to expend the time and energy necessary in completing the lengthy questionnaire, as well as their enthusiasm for understanding the sociological aspects of their research community is deeply appreciated. We are also grateful to Professors Thomas Allen and Lotte Bailyn for their comments on initial drafts of the survey instrument, and to Professors Edward Roberts and Roland Van Dierdonck for their generous financial support and encouragement.

For further information, please contact: Prof. M.A. Rappa, Building E51-338, 50 Memorial Dr., Cambridge, MA 02139, U.S.A. (Tel. 617 253-3627, Fax 617 253-2660)

Abstract

This report describes the preliminary results of an international survey of the neural network research community conducted over a three-month period between February 16 and May 11, 1990. The objective of the study is to examine the sociological dimensions of a research community, including identification and characterization of the phases that may occur over time as researchers attempt to develop a radically new technology. The scope of this report is limited to providing the reader with a brief overview of the rationale for the study and some of the basic statistics regarding the demographics of the survey respondents. Over the next six months additional reports will be produced describing the complete findings of the statistical analysis.

Introduction

Neural networks are a very interesting method of computation that is quite different in its approach from conventional rule-based techniques. Inspired by theories of neurological processes in biological organisms in their design, neural networks are typically implemented in areas such as machine vision, speech recognition, and signal processing--areas in which traditional computational methods have proven difficult. Their development has an intellectual history that spans more than forty-years. In recent years, the level of interest among researchers within several disciplines has risen dramatically, and numerous streams of activity have resulted in the commercialization of neural network computer software and hardware.

This study of neural networks seeks to understand the internal functioning of research communities and their role in contributing to the development of a radically new technology. It is part of a larger research program, begun in 1986, which currently includes investigations of twelve research communities. The results of these earlier studies indicate that research communities exhibit certain structural and behavioral characteristics, which may be important in understanding the underlying rate of scientific and technical progress. Specifically, these studies suggest a pattern of evolution within research communities reminiscent of three distinct phases, as illustrated in Figure 1.

The phases are diagramed in Figure 1 in terms of changes in the level of effort within the field. The first phase, which is coined the "bootlegging" phase, is characterized by a relatively low level of effort. During this period, a handful of researchers dedicate themselves to furthering the field, even though their enthusiasm may not be shared by their peers, and indeed, may be severely criticized. Typically, they have difficulty securing

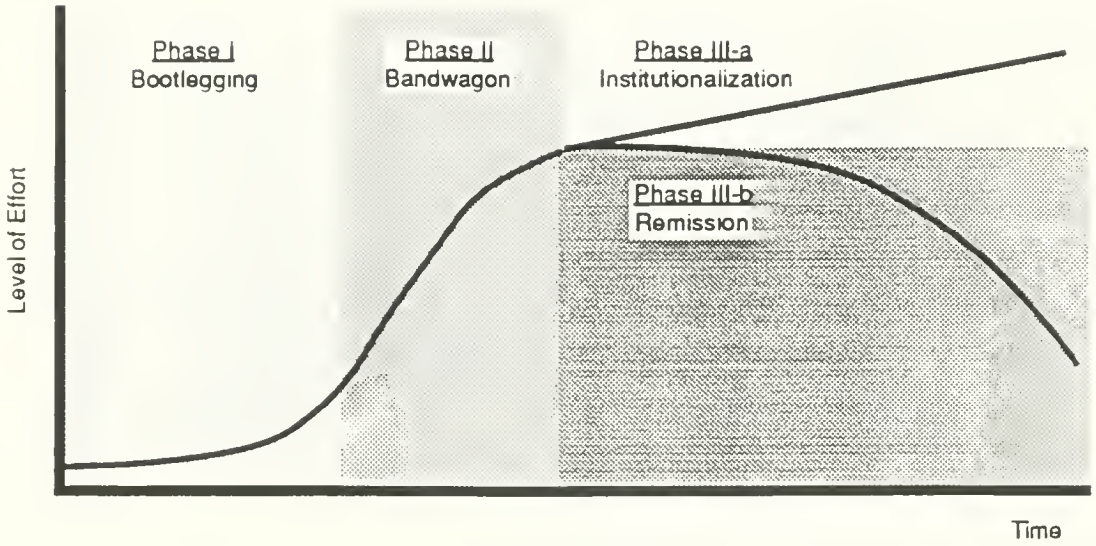

FIGURE 1: PHASE MODEL OF R&D COMMUNITY

adequate funding (hence the name bootlegging, which implies that researchers struggle to maintain their research without formal recognition or funding to underwrite the cost of their work.) Furthermore, there is no obvious market for the emerging technology. While the scientific value of the new field may be the topic of debate, the commercial value is extremely difficult to see at all. Thus, researchers tend to have little pecuniary motivation in pursuing their line of work.

In comparison, the second phase represents a sharp transition from the first phase. The dominant charactistic of this phase is the very rapid increase in the number of researchers working in the community in a relatively short period of time. This phenomenon is sometimes referred to as a "bandwagon;" hence the name for this phase. As the research community grows it typically becomes more widely distributed across organizations, sectors, and countries. Within this phase, the field begins to make headway against its mainstream detractors and builds wider scientific legitimacy. This second phase is also accompanied by the growth of niche markets, which begin to take advantage of the commercial potential of the new technology.

As the bandwagon progresses, the community enters the third phase, where one of two paths will emerge: (a) researchers continue to make progress in solving the problems confronting them, allowing the community to institutionalize itself, or (b) progress begins to

slow down such that researchers become discouraged, forcing the community to contract and perhaps eventually return to the conditions prevailing in the first phase.

Under scenario (a), success is at hand: the scientific and commercial value of the technology becomes increasingly apparent, markets begin to expand, funding sources are secured, and the community begins to institutionalize itself. The growth rate in the community participation will begin to stabilize at a lower level. Recruitment becomes more formalized as universities develop standard curricula and graduate programs, which allow students to specialize in the field. New university research centers dedicated to the field and new firms that specialize in the development and application of the technology are formed. Moreover, specialized journals, conferences, and professional societies that facilitate communication among researchers are established. Standards begins to emerge, easing the way for market development. In sum, what had once been a fringe group begins to establish itself within the academic and industrial mainstream.

The alternative scenario (b), illustrates a community in which progress does not meet researchers' expectations, markets do not develop, and the institutionalization process does not take hold. Under these circumstances, the large majority of researchers begin to filter out of the community in favor of other more promising research topics. Recruitment of new researchers suffers, and funding sources begin to run dry. The result is for the community to revert back to phase-one conditions. This implies that over longer periods of time, communities may undergo several cyclical fluctuations before successfully institutionalizing itself.

The Neural Network Community Survey

The objective of the neural network community survey is to probe more deeply into the sociological dimensions of the different phases and the factors precipitating the phase transitions. In particular, the survey focuses on Phases I and II in an attempt to better understand the attitudes and behavior of researchers who first enter the field in each phase: that is, to understand the similarities and differences which may exist between "bootleggers" and "bandwagoners."

The survey instrument is a 12-page questionnaire designed specifically for collecting information from neural network researchers about their research activities. The questions in the survey instrument are organized into five sections that deal with: (1) the extent and nature of the respondent's activities with neural networks; (2) the respondent's decision to work in the

field of neural networks; (3) the extent and nature of the respondent's communal activities, such as sharing information with other researchers through conferences, papers, and other collaborations; (4) the respondent's perceptions regarding the future progress of neural networks research; and (5) the basic demographic characteristics of the repondent and his or her organization.

The planning and design stage of the survey was conducted from November 1989 to February 1990. During this time a mailing list of the target population for the survey was constructed using authorship in scientific and technical journals, conference proceedings, workshops, and summer schools dedicated to neural network research over a two-year period from 1988 to 1989 (fourteen different literature sources in total). The target population numbered 2,037 individual researchers from thirty-five countries.

The questionnaires were mailed on February 16, 1990. A follow-up letter was mailed on March 1. Questionnaires were received up until May 11, 1990, at which time the last of the data were entered into the data file and a system file (using SPSS) was created to facilitate statistical analysis.

Preliminary Results

Over the three-month data collection period, a total of 710 individuals responded to the survey (see Figure 2). This represents an effective rate of return of 38%, which is considered satisfactory for an international survey of this size and scope. The return rate must be judged in light of the fact that budget constraints did not allow for the survey (which was in English) to be translated into the native languages of the researchers. The majority of the researchers in the target population lived in English-speaking countries, and many of those who did not, have published in English language journals. Moreover, because mailing addresses for the target population were derived from documents published in 1988-89, the problem of invalid addresses due to employment changes was fairly common. Other factors contributing to the return rate are: the length of questionnaire, which is long by most standards; coding of the questionnaires, which was necessary in order to provide the results of the study to participants; and the inability to provide return postage due to the costs involved and the international scope of the survey (although, return postage was guaranteed). All things considered, the 710 completed questionnaires is sufficient in number for most, if not all, of the statistical techniques to be employed.

Figures 3 and 4 illustrate the regional distribution of respondents and the number of respondents by country (for the top-10), respectively. The regional distribution of respondents corresponds very closely to that of the sampled population. Most respondents are working in the U.S. (60%); about one-quarter are based in Europe and 10% are based in the Far East. Although it is plausible that the survey is skewed toward U.S.-based researchers, an examination of the scientific literature on neural networks in 1988-89 shows a roughly similar regional distribution of researchers.

The sectoral distribution of respondents is shown in Figure 5. Slightly less than two-thirds of the respondents are employed in universities, while one-quarter are industrial researchers, with the balance in government organizations. Here again, the distribution corresponds to the target population and to the scientific literature on neural networks in 1988-89. Although not shown, there are a total of 216 universities represented among the respondents, 101 industrial firms, and 51 government organizations.

Figure 6 provides the disciplinary distribution of respondents, which shows that electrical engineers compose more than one-third of the total. When added together with computer scientists, the combined total for EE and CS is more than one-half of all respondents. The second largest category is the physical sciences, where physicists make up the majority of respondents. As with the sectoral and regional distributions, an analysis of the scientific literature on neural networks in 1988-89 shows a similar disciplinary distribution.

In terms of their position, 37% of the respondents hold faculty appointments (see Figure 7). The second most common position is that of staff scientist (or researcher), which represents about 33% of all respondents. This is followed by students (17%), engineers (7%), and managers (5%). Although not illustrated, the average age of respondents is 37 years and the median age is 35 years. The youngest quartile is between the ages 22 and 30; the oldest quartile is between ages 43 and 69.

Figure 8 shows the current funding for neural network research broken down by source for the respondents. The largest source of funds for respondents are their employer (averaging about 42% of total funds), followed by government agencies (averaging about 39% of total funds). Slightly more than one-third of government funding is defense-related. Personal funds, industrial sponsors, and private foundations make-up the balance of funding.

In terms of their level of interest in neural networks, Figure 9 shows that over 50% of the respondents view neural networks as their major or only research interest. In contrast, about 40% view it as one of many research interests. Only about 10% see it as a minor issue in their research agenda. Figures 10 and 11 provide a breakdown of the theoretical interests and various application areas of respondents. Using the categories set-out in the DARPA Neural Network Study (1988), respondents were asked to indicate the theory and application areas on which their research focuses. Simulation is the largest theory category, followed by learning and algorithms. Classification is the largest application category, followed by vision and signal processing. Although not shown, 40% of the respondents indicated that their neural network activities are to a great extent "basic research" in nature; 10% indicated "oriented basic research" as most characteristic of their work; 17% identified "applied research"; and 11% claimed that their activities were to a great extent "development" in nature. About 19% of the respondents hold one or more patents related to neural networks.

Figure 12 shows the number of respondents entering into the neural network field in each year, from 1950 to 1990. The data indicate that entry reached a peak in 1986 and 1987, with a total of 150 and 160, respectively. However, the decline in entry in the most recent years is explained in part by the techniques used to identify the survey sample population and not necessarily by an actual decline in the entry of new researchers.

In order to identify the first two phases of the community's evolution, a graph is made of the cumulative number of respondents over time as indicated by their year of entry (see figure 13). The graph shows both the cumulative number of respondents in the field (on the left axis) and cumulative percent (on the right axis). It clearly illustrates the rapid growth in the field, which began about 1984. About 25% of the respondents entered the field prior to 1984, whereas about 75% entered from 1984 to 1990. In accordance with the phase model presented in the introduction, in the subsequent analysis the period from 1950-1983 will be designated as Phase I, or the "bootlegging" phase; and the period from 1984-1990 will be designated as Phase II, or the "bandwagon" phase.

A major point of interest is in understanding the similarities and differences in researchers who enter the community during the bootlegging phase in comparison to those who enter during the bandwagon phase. The preliminary analysis suggests that there are several significant differences. First, an investigation of respondents' entry into neural network research relative to their graduate education is shown in Figure 14, using a scatter-plot of respondents' year of

entry into the neural network field by their year of graduation (for highest degree). A forty-five degree line through the origin highlights the distinction between researchers who started working on neural networks before graduating (i.e., they entered the field as students), and those who started after graduation. Also, a horizontal line through the year 1983 segments the data into the two phases of interest. Thus, quadrants 2 and 3 include respondents who entered the field after graduating, with 2 containing respondents who started work in the bootlegging phase and 3 containing respondents who started in the bandwagon phase. Similarly, quadrants 1 and 4 include respondents who started neural network research as students, with 1 containing students in the bootlegging phase and 4 containing students in the bandwagon phase. As can be seen more clearly in Figure 15, the bar diagram illustrates that among respondents in the bootlegging phase, the large majority started in the field as students. This contrasts markedly with the bandwagon phase, where the large majority of respondents entered the field after graduation.

A preliminary analysis was also made of the factors influencing respondents' entry into the field, as well as those factors that might influence respondents' exit from the field. The respondents were asked to indicate the importance of a variety of factors on their decision to work in the field of neural networks, as well as the importance certain factors might have on diminishing their interest in continuing to work in the field. Again, the responses were analyzed according to the phase in which the respondent began working in the field. The bar diagrams in Figure 16 and 17 indicate the mean value for each group for each factor in the entry and exit decision, and whether or not the difference between groups is statistically significant.

In terms of the entry decision (Figure 16), bootleggers were more influenced by the "intellectually compelling nature of neural networks" than their bandwagon counterparts; they were less influenced by the "positive opinions of other researchers"; they were less influenced by the "recent successes of other neural network researchers"; and they were less influenced by the "potential for peer recognition" than their counterparts in the bandwagon phase. There is no significant difference in the influence of "dissatisfaction with the previous research agenda" or the potential for "solving an important societal problem" between the two groups. However, bootleggers were less influenced in entering the field by the "availability for funding" of neural network research than bandwagon respondents; they were less influenced by the "lack of other promising research topics"; they were less influenced by the "potential for financial rewards"; and they were less influenced by the "opportunity to build a commercial enterprise" than their bandwagon counterparts.

As with the entry decision, there are also several statistically significant differences between the two groups in terms of the factors which might influence the decision to diminish respondents' commitment to neural networks (see Figure 17). The only category in which there is no significant difference between the two groups is in terms of a "diminished intellectual interestingness" of neural networks. However, bootleggers are less likely to be influenced by "a lack of funding for their research" than their bandwagon counterparts; they are less likely to be influenced by "increased financial costs of conducting neural network research"; they are less likely to be influenced by "rapid progress in alternative areas of research"; they are less likely to be influenced by "overcrowding in terms of the number of neural network researchers"; they are less likely to be influenced by "difficulty in keeping up with new development in neural networks"; they are less likely to be influenced by "slow progress in solving technical problems in neural networks"; they are less likely to be influenced by "discontinuance of neural network activities at their organization"; they are less likely to be influenced by "a lack of financial rewards"; they are less likely to be influenced by "diminished interest among other researchers in neural networks"; they are less likely to be influenced by "negative opinions of supervisors"; and lastly, bootlegger respondents are less likely to be influenced by "opinions of leading researchers unfavorable to neural networks" than their bandwagon counterparts.

Lastly, respondents were asked to indicate the number of years they are willing to work on neural networks before reaching their goals. Figure 18 shows the comparison between responses for researchers entering in the two phases, in terms of the percent of respondents in each category within each group. The data show that 50% of the respondents entering in the bootlegging phase are willing to work for 10 or more years in the field before reaching their goals. In comparison, the majority (about 46%) of researchers entering in the bandwagon phase are willing to work for as long as 2-5 years before reaching their goals.

Concluding Observations

The preliminary results of the international neural network survey suggest that the return rate is sufficient and that the sample corresponds to many of the characteristics of the target population. The first stage of analysis conducted so far is encouraging in that it shows evidence of distinct differences between the initial two phases in the proposed model of the evolutionary development of research communities. Nevertheless, more conclusive evidence will require further analysis.

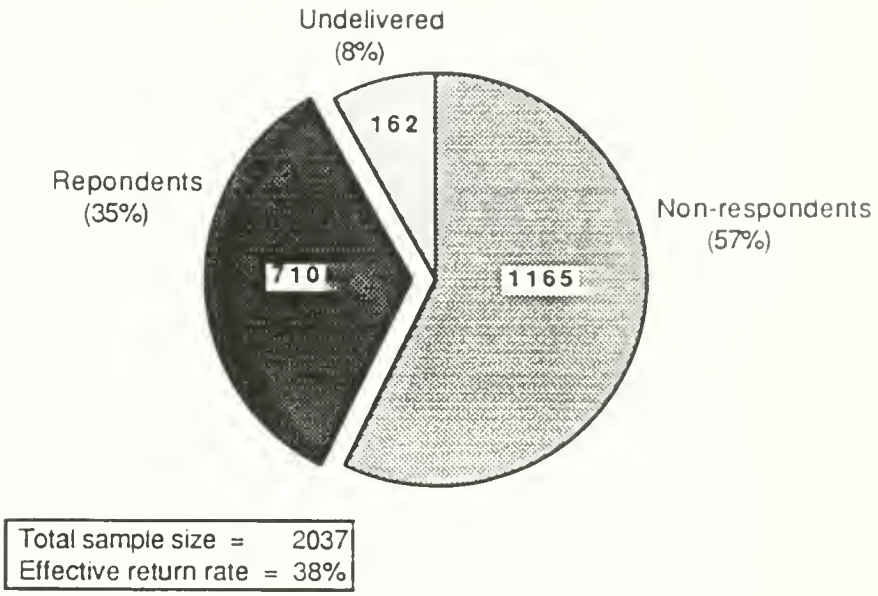

FIGURE 2: NEURAL NETWORK SURVEY RESPONSE RATE

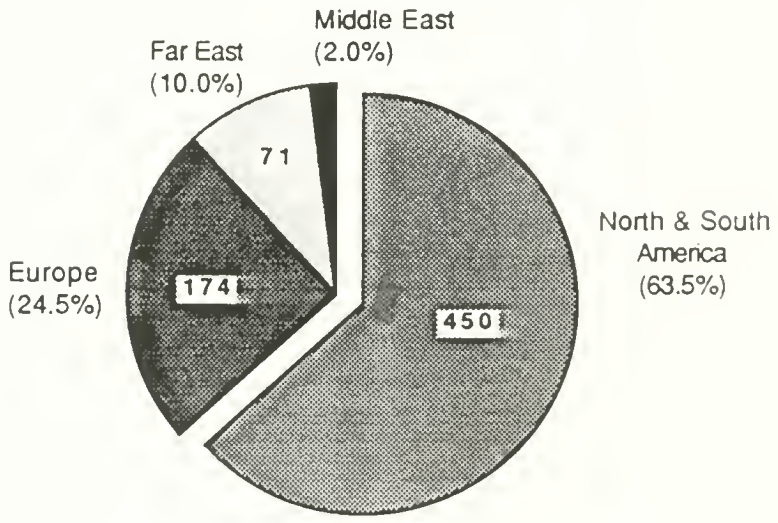

FIGURE 3: REGIONAL DISTRIBUTION OF RESPONDENTS

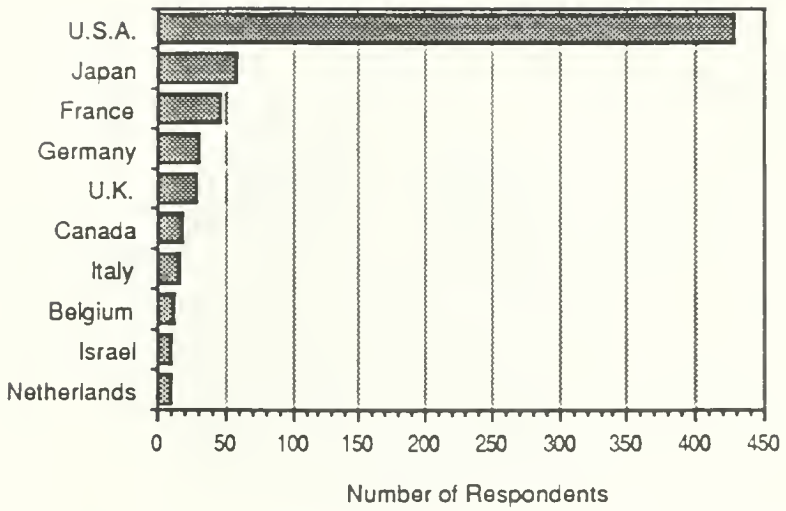

FIGURE 4: NUMBER OF RESPONDENTS BY COUNTRY (TOP-10)

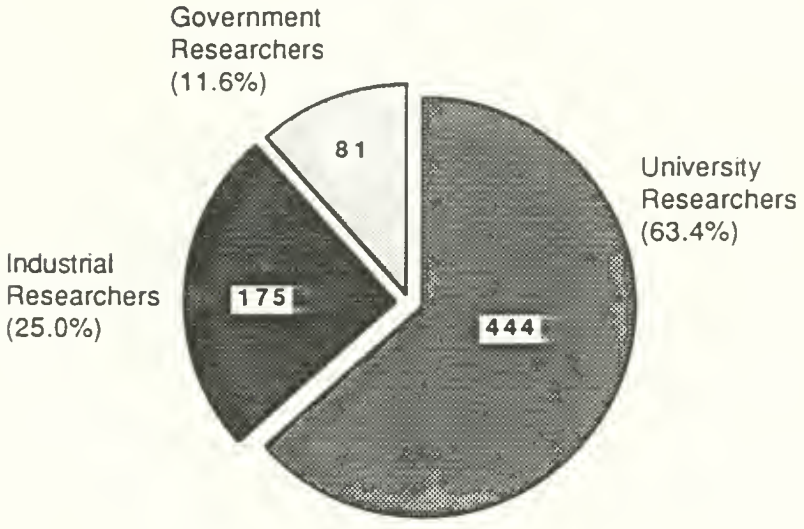

FIGURE 5: SECTORAL DISTRIBUTION OF RESPONDENTS

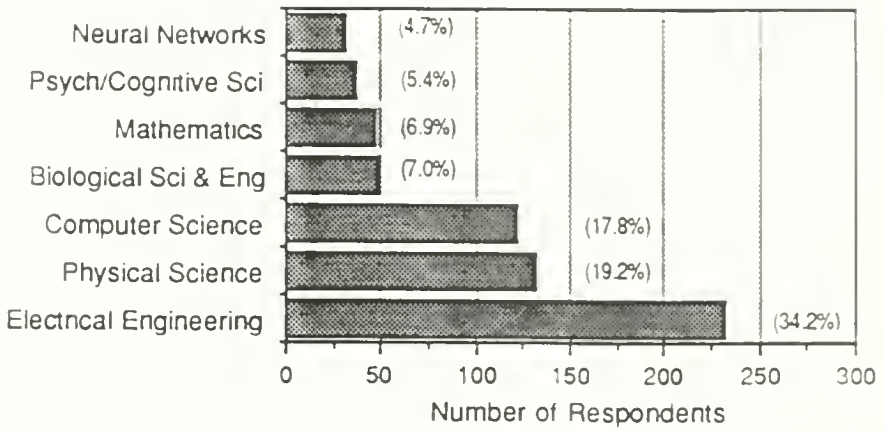

FIGURE 6: DISCIPLINARY DISTRIBUTION OF RESPONDENTS

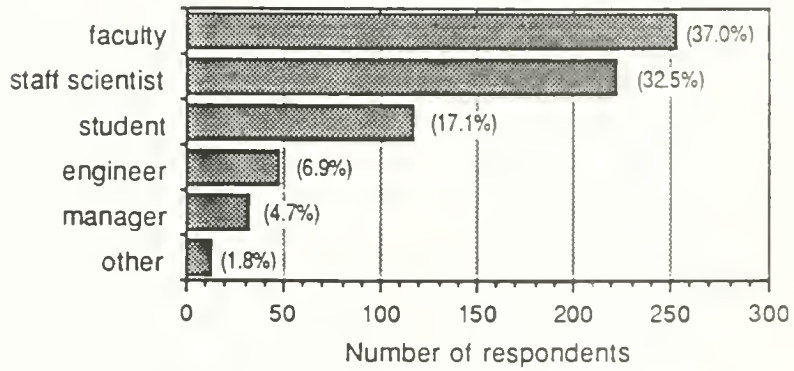

FIGURE 7: RESPONDENTS' CURRENT POSITION OF EMPLOYMENT

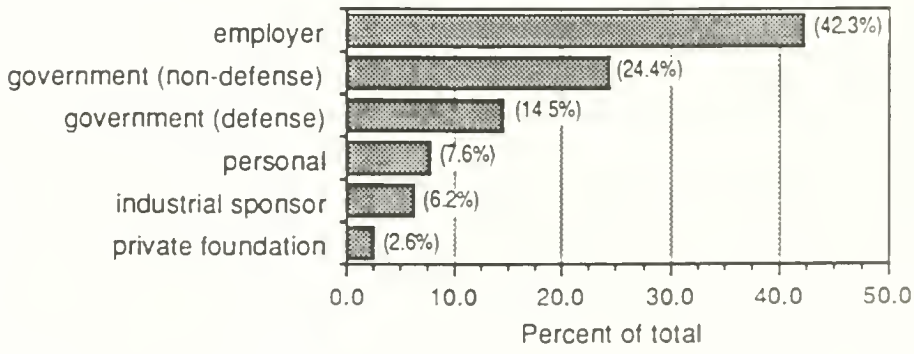

FIGURE 8: PERCENT OF RESPONDENTS' CURRENT FUNDING BY SOURCE

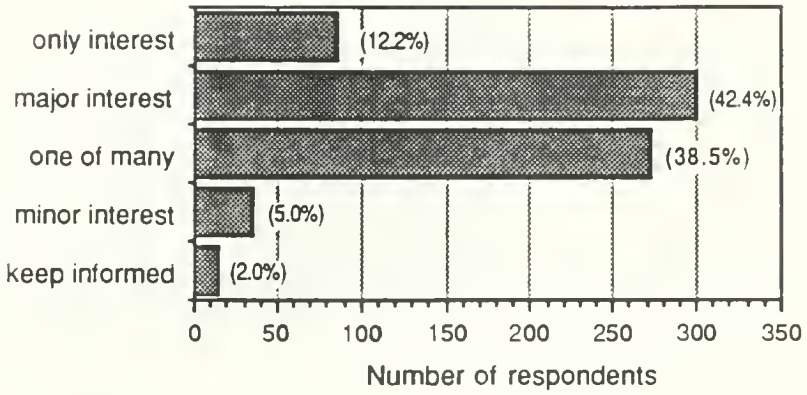

FIGURE 9: RESPONDENTS' LEVEL OF INTEREST IN NEURAL NETWORKS

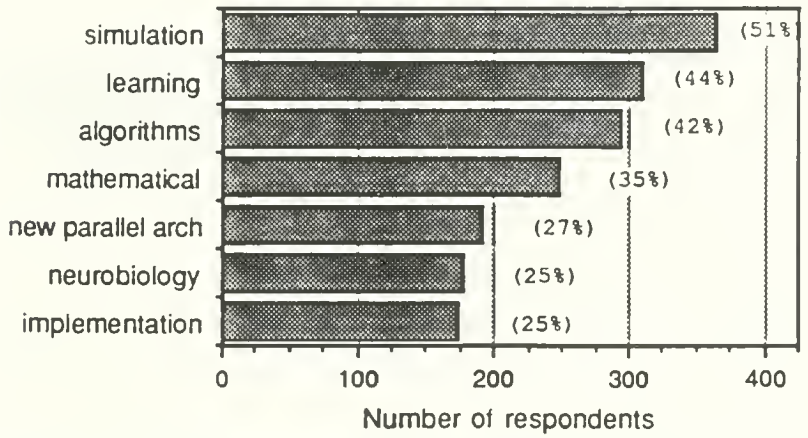

FIGURE 10: RESPONDENTS' INVOLVEMENT IN THEORETICAL ASPECTS OF NEURAL NETWORKS

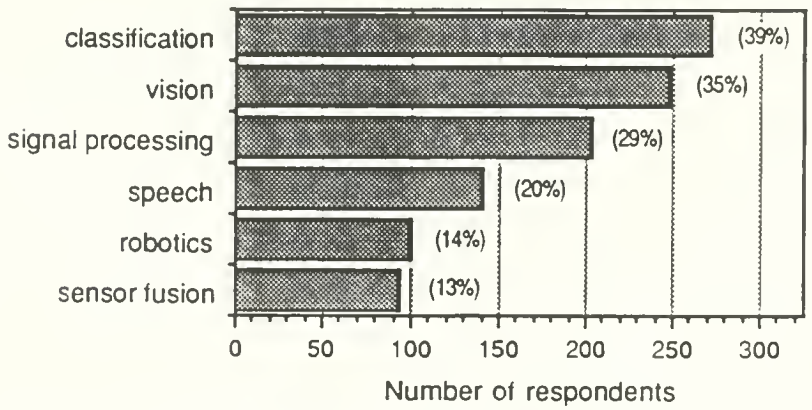

FIGURE 11: RESPONDENTS' INVOLVEMENT IN APPLICATIONS OF NEURAL NETWORKS

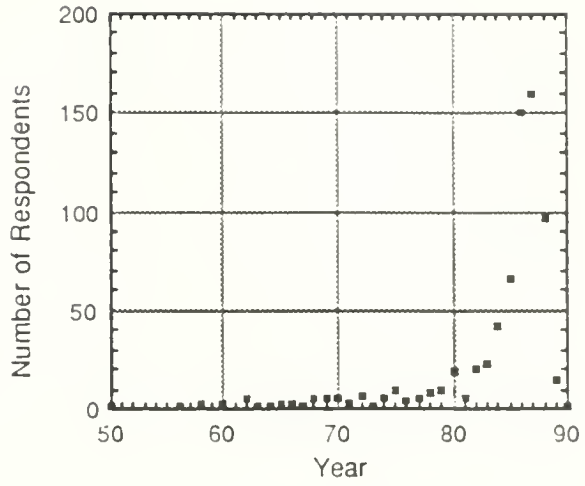

FIGURE 12: RESPONDENTS' YEAR OF ENTRY INTO THE FIELD OF NEURAL NETWORKS

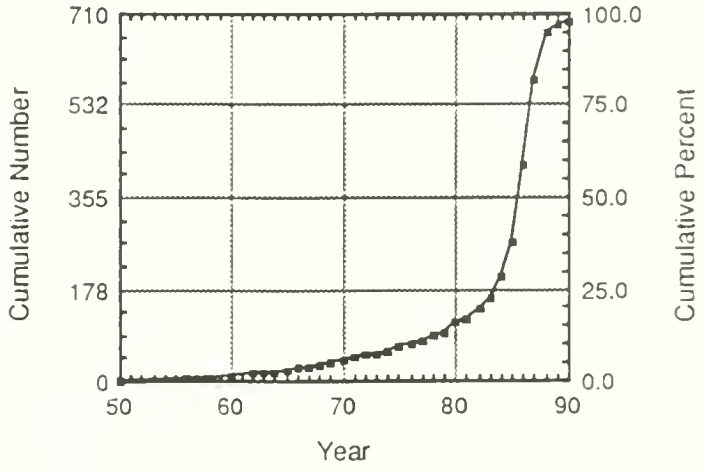

FIGURE 13: CUMULATIVE NUMBER OF RESPONDENTS IN THE FIELD AND CUMULATIVE PERCENT

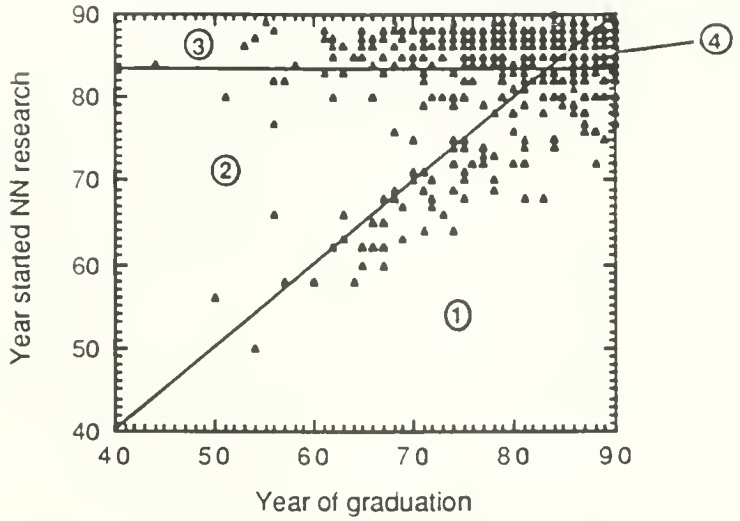

FIGURE 14: RESPONDENTS' FIRST YEAR OF NN RESEARCH RELATIVE TO YEAR OF GRADUATION (HIGHEST DEGREE)

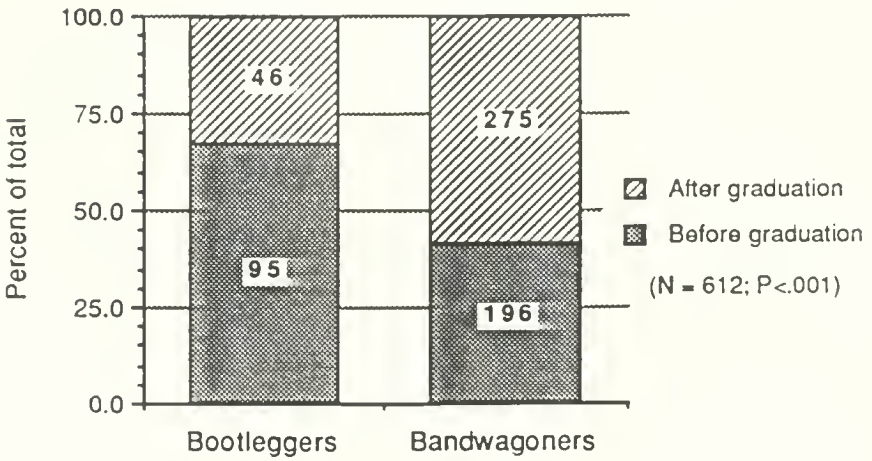

FIGURE 15: RESPONDENTS' ENTRY INTO FIELD RELATIVE TO GRADUATION

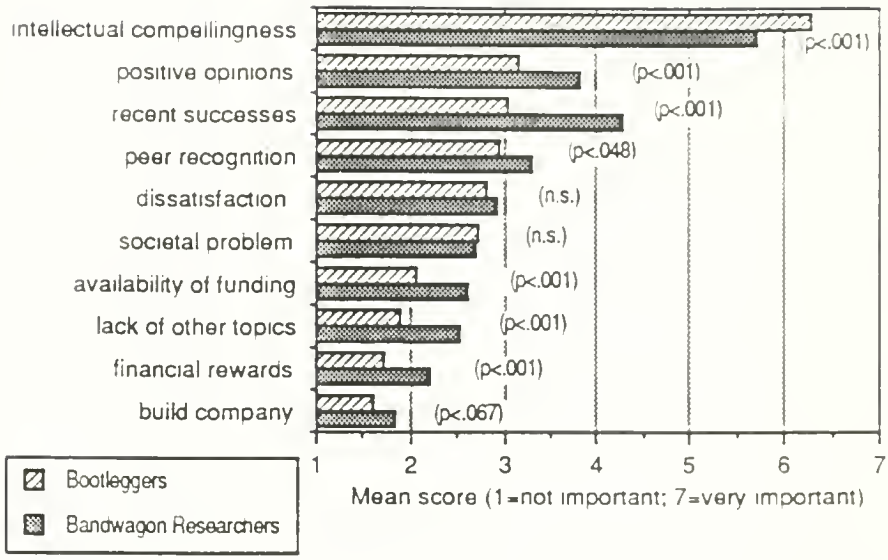

FIGURE 16: FACTORS INFLUENCING RESPONDENTS' ENTRY INTO THE FIELD OF NEURAL NETWORKS

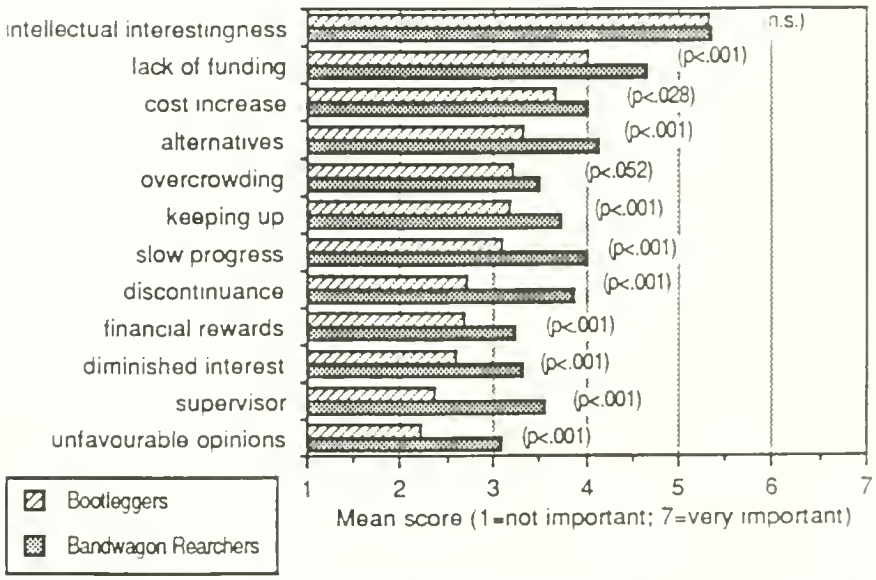

FIGURE 17: FACTORS POTENTIALLY INFLUENCING RESPONDENTS' EXIT FROM THE FIELD OF NEURAL NETWORKS

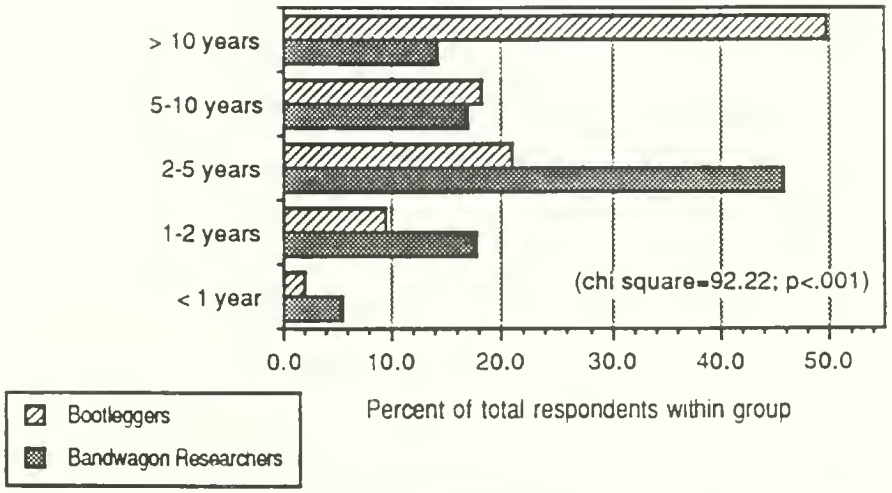

FIGURE 18: NUMBER OF YEARS RESPONDENT IS WILLING TO WORK ON NEURAL NETWORKS BEFORE REACHING GOALS